Small Patient Condition Stream Generator (SPCSG): Version 2.0

Wesley Sherman
Gerry Pang

Naval Health Research Center

Document No. 11-4D

Naval Health Research Center
140 Sylvester Rd.
San Diego, California 92106-3521

Small Patient Condition Stream Generator (SPCSG)

Version 2.0

Wesley Sherman
Gerry Pang

Naval Health Research Center
140 Sylvester Rd.
San Diego, CA 92106-3521

Report No. 11-XX, supported by the U.S. Air Force Medical Logistics Agency, under Work Unit No. 60334. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Navy, Department of Defense, or the U.S. Government. Approved for public release; distribution is unlimited. This research has been conducted in compliance with all applicable federal regulations governing the protection of human subjects. No human subjects were directly involved in this research.

Table of Contents

Executive Summary

The Small Patient Condition Stream Generator (SPCSG), developed by Naval Health Research Center (NHRC), allows NHRC personnel to make appropriate supply and supply quantity recommendations for Authorized Medical Allowance Lists. A tool was needed that would accept a small number of patients and still produce a valid estimate list of patient conditions and their associated probabilities. As a result, the SPCSG was developed, and its output has enabled further analysis of limited patient encounters. Other programs are available for estimating large patient streams, however, they do not provide accurate results for especially small ones, like those encountered in Mobile Forward Surgical Teams, Forward Resuscitative Surgical Systems, and Critical Care Air Transport Teams. This document describes the SPCSG, details the program's underlying methodology, and provides user instructions.

Introduction

The Small Patient Condition Stream Generator (SPCSG), developed by Naval Health Research Center (NHRC), helps NHRC personnel make appropriate supply and supply quantity recommendations for Authorized Medical Allowance Lists (AMALs). Analysts occasionally need to study scenarios where very small patient streams are distributed over a large number of available patient conditions (PCs). While other patient stream analysis tools are available, they are less effective when used for patient streams less than 100. The SPCSG allows analysts to precisely estimate supplies for these small patient streams.

AMALs and patient stream estimations are integral to NHRC's Estimating Supplies Program (ESP). Running a number of patients through the SPCSG over many iterations allows the user to determine which PCs are likely to occur and what supplies are required to appropriately treat those patients. The SPCSG may be used, for example, to determine a likely PC distribution for those functional areas that expect a small number of patients (less than 100), such as Mobile Forward Surgical Teams, Forward Resuscitative Surgical Systems, and Critical Care Air Transport Teams.

Methodology

The SPCSG applies a Poisson distribution algorithm to a list of PCs and their probabilities. The outputs, based on a small number of patients (less than 100) and other criteria selected by the user, are randomly generated patient streams for each user-entered iteration and summary statistics. An iteration from the detail output file may then be imported into a number of programs, including NHRC's Expeditionary Medical Knowledge Warehouse (EMedKW), Tactical Medical Logistics (TML+) planning tool, or ESP, to model and simulate medical scenarios that have small groups of patient encounters. The patient stream represents the rarity of injury and illness types that may occur, taken *n* at a time, where *n* is the initial, small number of patients entered by the user. Running multiple iterations "stabilizes" the results by increasing the likelihood that PCs with higher probabilities will have patients assigned to them over all the iterations. (The SPCSG data dictionary [see Appendix A] lists the tables and fields used in this application, including descriptions, types, sizes, and values for each field.)

Excel Source-File Requirements

The SPCSG accepts inputs from a Microsoft Excel spreadsheet (either an *.xls* or *.xlsx* file), a comma separated value file (*.csv*), or a text file (*.txt*). The input file must contain two data columns. The first is the list of PCs with a string data type, which will allow leading zeros such as "001;" the second column is the list of the probabilities for the associated PC. The source file probabilities must be written in decimal format and normalized (i.e., the sum of all PCs must equal 1), otherwise the file will be rejected and a warning message will display. In addition, any sized subset of PCs can be used, however small, but all probabilities should add up to 100%.

The application will overwrite any existing output files with the same name at the same location. If the user is interested in retaining previous results, the output files must either be renamed or moved to a different directory before the application is run again.

Small Patient Condition Stream Generator Form

The SPCSG uses standard Windows command buttons. These controls include the Minimize, Maximize, and Close buttons shown in the right corner of the form title bar in Figure 1.

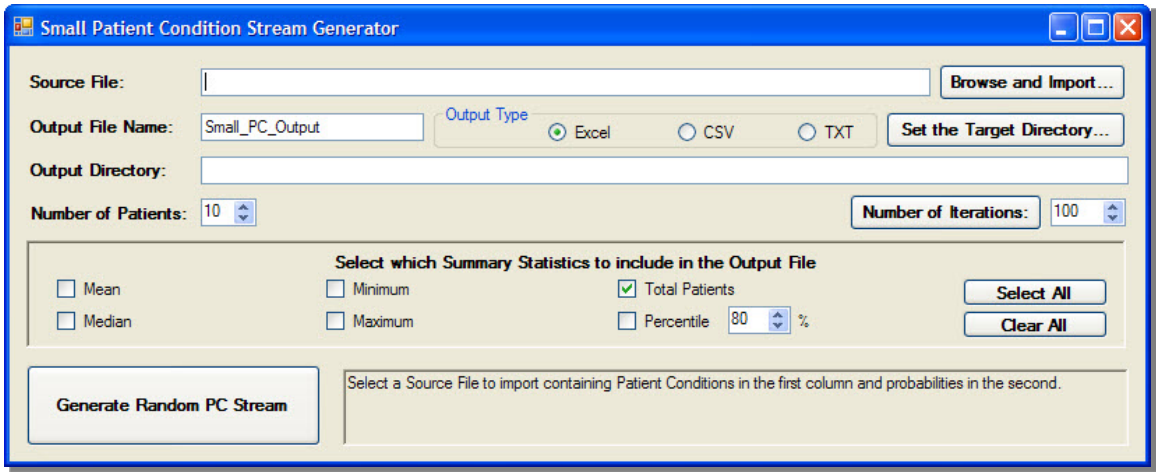

Figure 1. The SPCSG form window.

Source File Import

The Source File textbox is the box at the top of Figure 1. The user can cut and paste, type, or Browse and Import the source file name here. The entry must include the full pathway, including the file-type suffix. Special characters—other than an underscore, dash, slash, backslash, or period—trigger an error message that appears on top of the form; application processing suspends until the user acknowledges the message by clicking OK.

The best way to select the source file is to click on the Browse and Import command button to bring up the Select a Source File to Import dialog box (see Figure 2). Use the Files of type drop-down list to select the type of file to import, then locate the source file with the Look in drop-down list. Navigate to the file, select it, and click Open. The Source File textbox will update with the file name and full pathway.

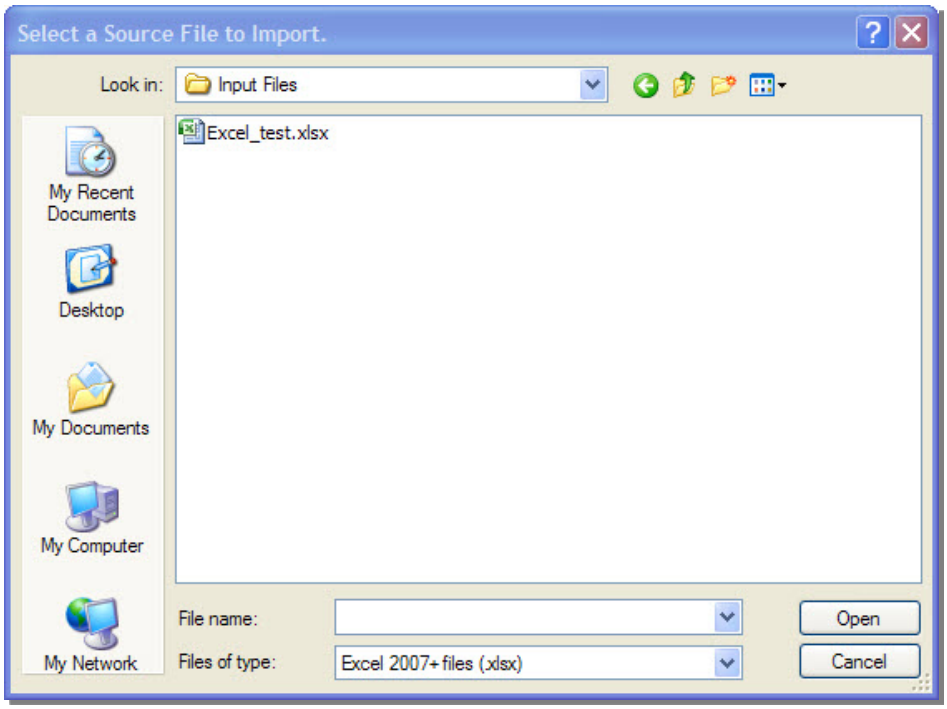

Figure 2. An example of the Select a Source File to Import dialog box.

If an Excel file is selected, the Input File Specifications dialog box appears. The user then selects the name of the worksheet to import, and whether the worksheet has a header row (see Figure 3). If a text file is selected, the Input File Specifications dialog box appears and the user selects the delimiter (tab or pipe), and whether the file has a header row (see Figure 4).

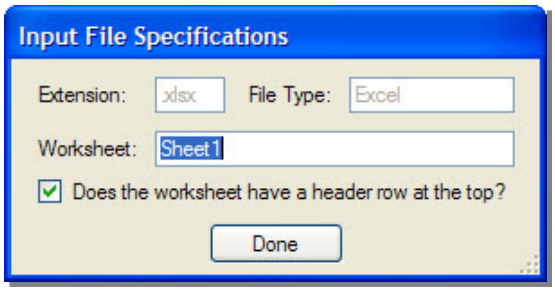

Figure 3. The spreadsheet file specification window.

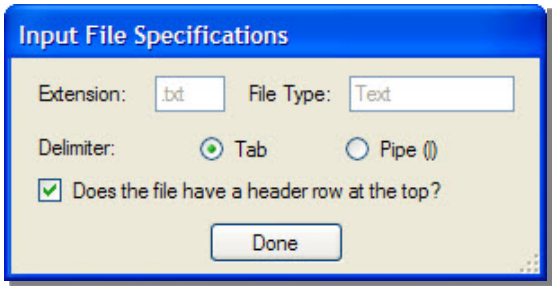

Figure 4. The text file specification window.

After selecting Done, the user is informed if the import of the source file was successful. If successful, the imported data is stored in the table PC_Stream.

Output File Controls

Having selected the input file above, the next step is to select the output file name. In the Output File Name dialogue box, the program automatically suggests the file name Small_PC_Output. At this point, the user can modify or replace the output file name. (Do not include a file type suffix; the suffix of the output file is set by the Output Type radio buttons).

To prevent overwriting previous output files, the user should move them out of the destination file, or ensure that the file names do not match. A matching file name in the Output File Name dialogue box will overwrite the previous file by the same name without warning.

Next, the user selects where the new file will be created by clicking the Set the Target Directory command button (see Figure 1). This brings up the Browse for Folder dialog box as shown in Figure 5. Navigate through the directory tree to select the target folder, and then click OK. The Output Directory textbox displays the selected directory, including its pathway. Once set, the mapped output directory will be the default location for all exported files until it is changed.

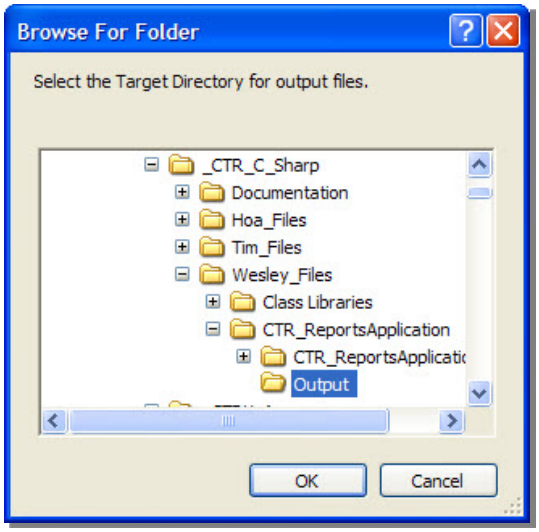

Figure 5. The Set Target Directory dialog box.

Algorithm Computations Controls

Next, in the SPCSG form window, the user enters the number of patients either by typing into the field or using the Number of Patients spinner (see Figure 1). The user may select a number from one to 100. This number passes to the Poisson algorithm for processing. The default value is currently set at 10.

The Number of Iterations spinner (far right) can be set the same way. This programs the number of times the set patient number is processed by the algorithm. The user may select any number from one to 10,000 (the default is 100). The number of iterations entered by the user will determine the number of records generated in the PC_Iterations table.

The user sets the percentile chance in the Get Percentile Chance dialog box (see Figure 6), which appears when the user clicks on the Number of Iterations command button. The percentile chance default value is 80%, but the percentile chance spinner can be set to any number from 1 to 100. Clicking Done closes the Get Percentile Chance dialog box. Now the SPCSG form window (see Figure 7) shows the percentile chance selected to the left of the Number of Iterations command button. This indicates the PC with the lowest probability in the bottom of the window (in this case, PC 98, with a probability of 0.000206516666072174), and updates the Number of Iterations spinner with the number of iterations required, according to the computations below.

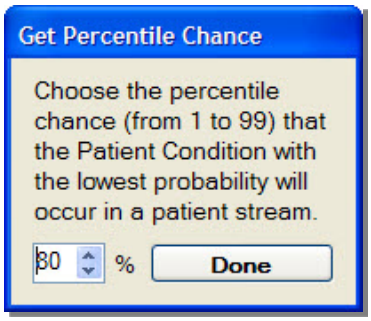

Figure 6. The Get Percentile Chance dialog box.

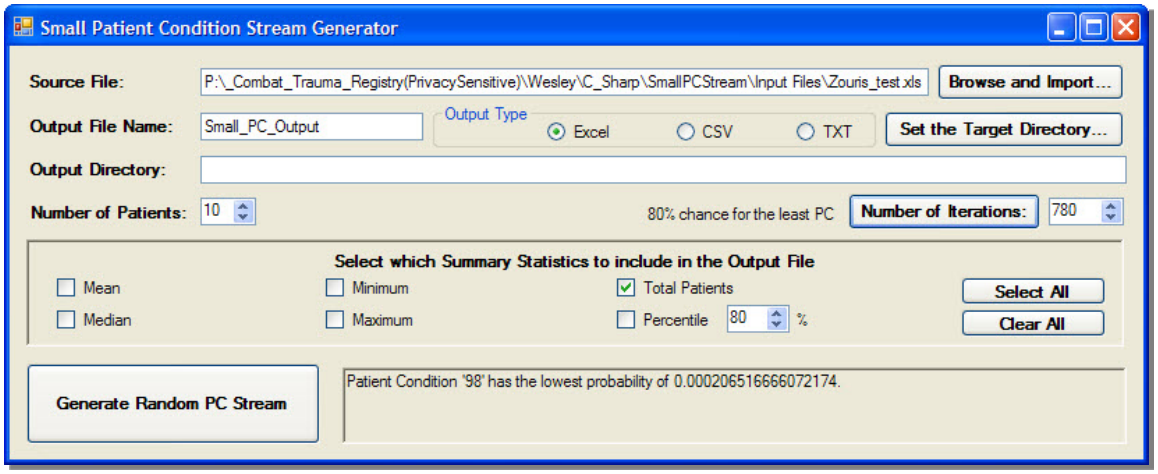

Figure 7. The SPCSG form after the Get Percentile Change dialog box closes.

The Number of Iterations calculation

The value of the Number of Iterations spinner (*n*Iterations) is automatically set using the following calculation:

$$\alpha = 1 - (n\text{Chance} \div 100) \quad (1)$$

$$n\text{Total} = n\text{Prob} \times n\text{NumPats}$$

$$\beta = 1 - e^{-n\text{Total}}$$

$$n\text{Iterations} = \text{INT}((\text{LN}(\alpha) \div (-\beta)) + 0.5)$$

Here, *n*Chance is the percentile chance set by the user in the Get Percentile Chance dialog, *n*Total is the estimated total number of patients, *n*Prob is the lowest probability in PC_Stream, *n*NumPats is the number the user set in the Number of Patients spinner, the INT() function takes the integer of the calculation, *e* represents the EXP() function for exponential computations, and the LN() function calculates the natural logarithm.

Using the example in Figure 7:

$$\alpha = 1 - (80 \div 100) = .2 \quad (2)$$

$$n\text{Total} = 0.000206516666072174 \text{ x } 10 = 0.00206516666072174$$

$$\beta = 1 - e^{-.0.00206516666072174} = 0.00206303567125532$$

$$n\text{Iterations} = \text{INT}((\text{LN}(.2) \div (-0.00206303567125532)) + 0.5) = 780$$

Finally, the SPCSG form window provides several summary statistics option controls (see Figure 7) within the box labeled Select which Summary Statistics to Include in the Output File. For each summary statistic check box, other than the Percentile check box, a column will be added to the Excel output file as that statistic is computed for each PC, in each iteration. The default selection is the Total Patients check box. At least one check box must be selected so that the program can produce an output. Summary statistics options include:

- The Mean check box produces the sum of all patients randomly generated over all iterations divided by the number of iterations (set in the Number of Iterations spinner), which results in a whole number.

- The Median check box determines the median by sorting the number patients generated over all iterations, in ascending order, and locates the figure for the middle record.

- The Minimum check box generates an integer that shows the lowest number of patients generated over all iterations (for most PCs, this will be zero).

- The Maximum check box displays the integer that shows the highest number of patients generated over all iterations.

- The Total Patients check box records the sum of the patients generated for all iterations (this is useful for running subsequent queries).

- The Percentile check box computes and stores the number of patients determined for the 65th to the 95th percentile records, sorted by the number of patients generated in ascending order in 5% increments. The user can set any percentile number from 50 to 99 by using the percentile spinner. Percentile figures for the 65th, 70th, 75th, 80th, 85th, 90th, or 95th percentiles will automatically be determined, and if the spinner is set to one of these seven numbers, nothing more will be done. However, if any other number is selected, additional figures will be determined for that percentile, and another column will be added to the output file. The Percentile spinner defaults to 80.

The Select All and Clear All command buttons are shortcuts that allow the user to rapidly check and uncheck all of the summary statistics check boxes and avoid toggling each one individually.

Having selected all processing parameters, the user begins the generator process by clicking Generate Random PC Stream (see Figure 7). The user can view progress and feedback in the adjacent read-only information box. This provides guidance to the user during the input phase.

Processing the User's Input

The following sections describe processing the user's input once the Generate Random PC Stream command button is selected. First, the imported probabilities in the PC_Stream table are summed. If the sum does not equal 1, then a warning is displayed in the information box, but processing continues.

Populating the PC_Iterations Detail Table

The estimated PC total is calculated for each PC in the PC_Stream table as follows:

$$n\text{Total} = n\text{Prob} \times n\text{NumPats} \tag{3}$$

Here, *n*Prob is the probability for that PC and *n*NumPats is the number of patients entered by the user. There are three possible outcomes:

1. Due to precision limitations, the Poisson algorithm cannot process estimated PC totals above 35. If a PC *n*Total is greater than 35, then an "overflow warning" appears in the information box listing the PC and its computed total. A single record is added to the PC_Iterations table with the number of patients listed as the negative of the computed number. Negative numbers in the PC_Iterations table are excluded when calculating the summary statistics in the PC_Stats table.

2. If the estimated PC total computes to 0, then a Zero Patients warning appears in the information box and a negative 1 is set for the number of patients for that PC in the PC_Iterations table, excluding this result from the summary statistics calculations.

3. If the estimated PC total passes both of the above checks, then it is processed through the Poisson algorithm and the randomly generated number of patients is added to the PC_Iterations table for each iteration. The information box keeps the user informed of progress as it updates every 100 iterations. Because the random Poisson algorithm is applied to each PC, the resulting totals for each iteration in the PC_Iterations table may not equal the number of patients entered by the user. For example, if the user entered 10 in the Number of Patients spinner, the randomly computed sum may be 7 for one iteration and 13 for another.

Populating the PC_Stats Summary Table

For each PC in the PC_Stream table, one or more statistic columns will be added to the PC_Stats table for each summary statistics check box that the user has checked. This means that there will be a row of statistics for each PC, with the statistics determined from the detail records in the PC_Iterations table. For example, if the Mean and Total Patients check boxes are checked, the numbers in the PC_Iterations table will be averaged and summed respectively for each PC, and columns for those statistics will be included in the PC_Stats table.

Negative numbers in the table PC_Iterations are simply copied to the PC_Stats table, since no calculations will be done with them. Otherwise, a statistic is determined for each summary statistics item selected by the user. These are added as a column for each PC record in the PC_Stats table.

A message box pops up and informs the user when the statistical computations for each PC are complete. When all processing is finished, any rejected PCs (those that were generated as negative numbers in the output tables) are listed.

After both the PC_Iterations and PC_Stats tables have been populated, they are exported (as specified) by the user in the output controls. The statistics file is exported with "_Stats" appended to the file name. If successful, the information box displays a confirmation; otherwise, an error box appears .

Poisson Algorithm

The algorithm that randomly generates the number of patients for each PC and each iteration does so in the following way. First, a multiplier (*n*Mult) is determined based on the expected number of patients (*n*Total) passed to the algorithm. The following *n*Mult values were determined to be the maximum number of times that the Poisson calculation will occur before it returns numbers that are too small:

*n*Total	*n*Mult
≤ 4	30
≤ 8	40
≤ 12	50
≤ 25	$4 \times n\text{Total}$
≤ 35	$3 \times n\text{Total}$

Second, the actual Poisson calculation is determined in a loop from 0 to *n*Mult, and the results saved in the Poisson table for each loop. The Poisson calculation is as follows:

$$\alpha = n\text{Total} \wedge n\text{Count} \text{ (i.e., } n\text{Total to the power of } n\text{Count)} \quad (4)$$

$$\beta = n\text{Count! (i.e., } n\text{Count factorial)}$$

$$n\text{Poisson} = n\text{Poisson} + e^{-n\text{Total}} \times \alpha \div \beta$$

Here, *n*Count is the loop count from 0 to *n*Mult, and *n*Poissonis initialized to 0 prior to the loop. A record is appended to the Poisson table as follows:

Field	Value
p_occ	*n*Count
p_prob	*n*Poisson

Third, the first record in the Poisson table with the highest value of *p_prob* less than 1 is located, the next record is set to 1, and all subsequent records are deleted. When completed, the records in the Poisson table compose a probability curve for a PC shown in the Figure 8 example.

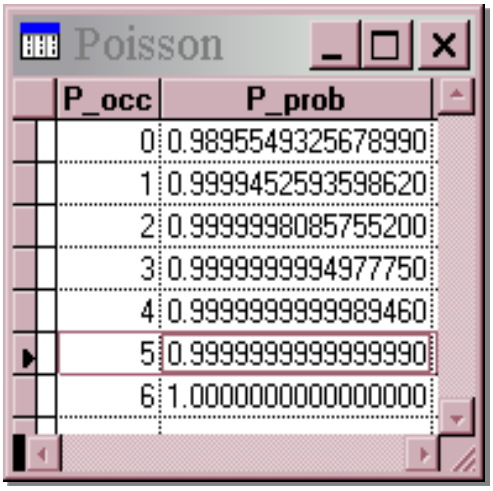

Figure 8. The resultant Poisson.dbf table.

Fourth, a random number between 0 and 1 is generated by the Random() function, and the record is located in the Poisson table with the minimum number of patients where:

$$Poisson.p\_prob \geq n\text{Rand} \tag{5}$$

Here, *n*Rand is the random number generated.

Finally, field P_occ in Figure 8 is the determined number of patients returned for that PC.

Conclusion

The SPCSG appears to provide greater accuracy for small patient stream estimations, as compared with the current large patient stream tools. Now it is possible to estimate valid probabilities for PCs with more confidence in situations and scenarios where small patient streams are encountered (as in Mobile Forward Surgical Teams, Forward Resuscitative Surgical Systems, and Critical Care Air Transport Teams). This output enables further analysis and precision.

Appendix A

Small PC Stream Generator Data Dictionary (May 2010)

PC_Iterations

Table Name: PC_Iterations **Database:** SmallPCStream

Description: Stores the number of patients randomly generated for each iteration of every PC. Emptied prior to each run.

Field Name	Type	Width	Description	Values
iterations	*smallint*		The iteration number.	Integer
pc	*varchar*	50	Patient Condition code.	PC_Stream.pc
num_pat	smallint		Randomly generated number of patients.	Integer
pat_order	smallint		Counter used to number the records for each PC in sequence.	Integer

PC_Stats

Table Name: PC_Stats **Database:** SmallPCStream

Description: Used to export the summary statistics. It has a column for each summary statistic chosen by the user. Emptied prior to each run.

Field Name	Type	Width	Description	Values
pc	varchar	50	Patient Condition code.	PC_Stream.pc
prob	*decimal*	26.25	Probability of a PC code expressed as a decimal, e.g., 50% is entered as 0.50000000000000000000000000.	PC_Stream.prob
pc_mean	decimal	27.24	Mean or average number of patients over all iterations. Optional field; at least one of the `pc_*` fields will be included in the table.	Integer
pc_median	smallint		The median or number of patients in the middle record of all iterations. Optional field.	Integer
pc_min	*tinyint*		The minimum number of patients determined for all iterations. Optional field.	Integer
pc_max	smallint		The maximum number of patients determined for all iterations. Optional field.	Integer

Table Name: PC_Stats **Database:** SmallPCStream

Description: Used to export the summary statistics. It has a column for each summary statistic chosen by the user. Emptied prior to each run.

Field Name	Type	Width	Description	Values
pc_total	smallint		The sum of the number of patients determined for all iterations. Optional field.	Integer
pc_median	smallint		The median or number of patients in the middle record of all iterations. Optional field.	Integer
percentile_user	smallint		The number of patients of the record just above the ## percentile record in PC_Iterations, where ## is the percentile chosen by the user.	Integer
percentile_65	smallint		The number of patients of the record just above the 65th percentile record in PC_Iterations.	Integer
percentile_70	smallint		The number of patients of the record just above the 70th percentile record in PC_Iterations.	Integer
percentile_75	smallint		The number of patients of the record just above the 75th percentile record in PC_Iterations.	Integer
percentile_80	smallint		The number of patients of the record just above the 80th percentile record in PC_Iterations.	Integer
percentile_85	smallint		The number of patients of the record just above the 85th percentile record in PC_Iterations.	Integer
percentile_90	smallint		The number of patients of the record just above the 90th percentile record in PC_Iterations.	Integer
percentile_95	smallint		The number of patients of the record just above the 95th percentile record in PC_Iterations.	Integer

PC_Stream

Table Name: Pc_Stream **Database:** SmallPCStream

Description: Holds the PC codes and their probabilities. The source table, which must contain only the columns matching the fields listed below, is converted into this table. The sorted PCs in this table are used to build the other PC_* tables. Emptied prior to each run.

Field Name	Type	Width	Description	Values
pc_text	varchar	50	Patient Condition code.	Text
prob	*numeric*	26.25	Probability of a PC code expressed as a decimal, e.g., 50% is entered as 0.50000000000000000000000000000000000.	Decimal number <= 1

Poisson

Table Name: Poisson **Database:** SmallPCStream

Description: Used in the calculation of the Poisson algorithm. Emptied prior to each run.

Field Name	Type	Width	Description	Values
p_occ	smallint		Loop count for determining the Poisson distribution.	Integer
p_prob	*float*		Poisson distribution maximum probability factor.	Decimal number

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD MM YY)
22 12 10

2. REPORT TYPE
Technical Document

3. DATES COVERED (from – to)
May 2005–Nov 2010

4. TITLE
Small Patient Condition Stream Generator (SPCSG): Version 2.0

5a. Contract Number:
5b. Grant Number:
5c. Program Element Number:
5d. Project Number:
5e. Task Number:
5f. Work Unit Number: 60334

6. AUTHORS
Sherman, Wesley; Pang, Gerry

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commanding Officer
Naval Health Research Center
140 Sylvester Rd
San Diego, CA 92106-3521

8. PERFORMING ORGANIZATION REPORT NUMBER
11-4D

8. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Commanding Officer
Naval Medical Research Center
503 Robert Grant Ave
Silver Spring, MD 20910-7500

Commander
Navy Medicine Support Command
P.O. Box 140
Jacksonville, FL 32212-0140

10. SPONSOR/MONITOR'S ACRONYM(S)
NMRC/NMSC

11. SPONSOR/MONITOR'S REPORT NUMBER(s)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Small Patient Condition Stream Generator (SPCSG), developed by Naval Health Research Center (NHRC), allows NHRC personnel to make appropriate supply and supply quantity recommendations for Authorized Medical Allowance Lists. A tool was needed that would accept a small number of patients and still produce a valid estimate list of patient conditions and their associated probabilities. As a result, the SPCSG was developed, and its output has enabled further analysis of limited patient encounters. Other programs are available for estimating large patient streams, however, they do not provide accurate results for especially small ones, like those encountered in Mobile Forward Surgical Teams, Forward Resuscitative Surgical Systems, and Critical Care Air Transport Teams. This document describes the SPCSG, details the program's underlying methodology, and provides user instructions.

15. SUBJECT TERMS
Small Patient Condition Stream Generator; SPCSG; small patient stream; patient condition; PC; Authorized Medical Allowance List; AMAL

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	18a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Commanding Officer
UNCL	UNCL	UNCL	UNCL	15	18b. TELEPHONE NUMBER (INCLUDING AREA CODE) COMM/DSN: (619) 553-8429